RAND

An Evaluation Strategy Developed by RAND for the Broad Foundation

.Tammi Chun, Gail Zellman, Brian Stecher, and Elizabeth Giddens

RAND Education

20011010 101

Preface

The Broad Foundation was founded in 1999 to promote the development of school system leadership as a means of improving the education of students in large, urban public school districts. In 2000, the Foundation approached RAND for assistance in designing an evaluation strategy that would help the Foundation monitor grant programs, assist grantees to improve their programs and build capability, provide compelling data about the value of promising approaches, and determine whether its overall funding strategy is meeting its own goals. Key in developing the evaluation strategy was the Foundation's wish to use evaluation resources efficiently, so that small local projects would not be unduly burdened, while larger, more ambitious projects would be adequately assessed. RAND staff worked closely with Broad Foundation staff to create an evaluation strategy that could be implemented in a relatively straightforward way and that could be modified over time to meet changing Foundation and grantee needs. A critical review of several current evaluation methods and an informal survey of other philanthropic organizations' evaluation efforts informed this effort.

The strategy that RAND developed, while tailored to the needs of the Broad Foundation, may prove useful to other foundations and funders of services in their efforts to monitor grantees and assess the efficacy of their overall funding strategies. This strategy will also further the discussion of evaluation approaches more generally.

The research was conducted within RAND's Education unit. The overall goal of this unit is to bring accurate data and objective analysis to important national education policy issues.

Table of Contents

List of Figures

List of Tables

"Making a Human Capital Bet on School System Leadership"

The Broad Foundation has a very clear goal: to develop "a new kind of school system leadership" that will improve large, urban public school systems so that students will reach higher standards (Broad Foundation, a). In a speech to the Council of Chief State School Officers, Dan Katzir, Broad's Director of Program Development, described the foundation's effort as "making a human capital bet on school system leadership." He explained that the foundation believes that strengthening school leadership will "set the conditions for improved student learning and therefore ultimately affect higher academic achievement results for all students" (Katzir, 2000). In short, the Broad Foundation believes that better leadership can serve as the catalyst to spark and support school progress of many kinds.

Consequently, the foundation directs its funding toward recruiting, developing, and helping leaders who play key roles in the realms of governance (school boards), management (superintendents and principals) and labor relations (teachers). It funds programs whose activities match one or more of five objectives (Katzir, 2000):

- Enlisting talented people to serve as leaders in urban schools/education

- Redefining the roles and authorities for governing bodies, managers and labor unions

- Building leadership capacity

- Providing incentives for results

- Honoring success.

To support these objectives, the foundation operates "as an entrepreneurial grant-making organization" that values innovative approaches to leadership needs and challenges (Broad Foundation, b).

In order to learn whether Broad programs are effective at achieving these objectives, the foundation needs a means of evaluating its grants. To investigate the best approach to evaluation for the foundation, RAND conducted interviews with Broad Foundation staff and researched the evaluation approaches currently in use by other foundations. See Appendix A for a summary of our research on the evaluation purposes, measures, processes, capacity and support for grantee program evaluations of other foundations. See Appendix B for descriptions of three notable evaluation approaches currently used by other foundations.

This report is designed to help the Broad Foundation assess the impact of its portfolio of grants by extending and improving its and its grantees' use of evaluation. RAND proposes an approach to evaluation that is uniquely suited to Broad's needs and operations. The strategy's primary purpose is to systematize foundation data collection through the development of reporting requirements and evaluation expectations for grantees. The benefits of instituting such a strategy are, we think, substantial. Once it is applied to the foundation's portfolio of grants, staff will be able to ensure program accountability, identify promising practices and lessons learned, collect evidence to support scaling-up efforts through policymaking and technical assistance, and, in some cases, measure Broad's impact on public discussion of educational issues.

Challenges in Evaluating Leadership Programs

The importance of educational leadership to good student outcomes is indisputable, but it is difficult to trace the causal chain between leadership and achievement. It seems obvious that talented, well-informed school leaders can have a significant influence on the achievement of a school or district. However, there is very little empirical evidence about how effective school leaders operate. In fact, no broadly accepted set of measurable outcomes exists to assure grantees or foundation staff that programs to promote strong and creative leadership are working as intended. Research shows that student achievement is influenced by a number of factors: family characteristics (Coleman, 1996), policies such as class size (Word et al., 1990), resource allocation, teacher qualifications (Ferguson, 1991), training (Cohen and Hill, 1998), and the leadership of entrepreneurial, "dynamic" principals (Gates, Ross, and Brewer, 2000). Identifying the unique contribution of good leadership is very difficult.

Because leadership development initiatives are yet another step removed from the classroom than leaders themselves, however, returns on investment in developing educational leadership are even more difficult to measure than the effects of leaders. Tracing a path backwards—from classrooms, school administration, the skills and status of individual leaders, the strategies they use, and the decisions they make—to specific program activities is possible only via an attenuated causal chain. In many cases, detecting links between leadership development programs and student achievement may largely be a matter of drawing reasonable, though not verifiable, inferences. A second difficulty of assessing leadership development programs is isolating the effects of one program from those of other events and initiatives. Particularly in the short run, one must be cautious in labeling causes and effects too readily. This situation presents a quandary because outcomes—improving performance at the child level (e.g., improved test scores)—are clearly the long-term "gold standard" for any educational improvement program.

Making the case that a leadership initiative affects student achievement is quite difficult, and yet, it is the claim one cares most about.

Furthermore, identifying the specific contributions of programs to student outcomes can be problematic. On the one hand, verifiable evaluations of programs often require intensive data collection and analysis, which can be costly and invariably take time. On the other, many programs do not warrant an intense level of scrutiny. For example, evaluators may be tempted to examine small interventions and their impact on large problems. Participants in a Brookings symposium on evaluation of social interventions cautioned about measuring the effectiveness of such interventions which are not "functioning at a scale and level of intensity to make it reasonable to believe they may be successful," and whose success depends on the occurrence of "other important, non-trivial changes in the system ... which the reform itself cannot bring about" (Brookings, 1998, 8). Gary Walker of Public/Private Ventures noted, "The probability of anything coming out of these efforts is so small that it would be a shame to waste resources evaluating them—and a shame to end up with another stack of evaluations that shows that nothing works" (Brookings, 1998, 13). Striking the chord of appropriate evaluation becomes, then, a matter of doing something for every project—evaluating each one in some way—but doing only as much as can be truly useful to the foundation or grantee.

In order to create a strategic evaluation strategy for Broad's portfolio of grant programs, RAND needed to identify a flexible, transparent system of effort- and cost-appropriate evaluation for programs of varying scope, duration, ambition, and impact. A flexible approach would enable foundation staff and grant recipients to tailor the evaluation to the project. A transparent system would assure that grantees, the foundation and evaluators are clear about the purposes of the evaluation, the anticipated evaluation measures, the theories connecting grant activities to outcomes, and when and how the grant-funded efforts are expected to yield the anticipated outcomes.

Given these goals, we propose the following principles as a basis for the evaluation strategy:

- Accountability: Grantees should be accountable to the Broad Foundation and demonstrate that funds are both well spent and advance the goals and purposes for which they were awarded by the foundation.

- Appropriateness: The scope and intensity of each project's evaluation should be commensurate with the role that the project plays in the foundation's overall strategy. The

resources allocated to the evaluation (e.g., staff time and other costs) should be determined by the magnitude of the project as well as by the grantee's capabilities.

- Partnership: Broad Foundation staff and grantees should share the responsibility for program evaluation and work together to execute them. In some cases, external evaluators may also join the team when their expertise or impartial perspective is essential.

- Capacity building: As with venture capital firms, the Broad Foundation invests in the ideas and leadership of its grantees. Evaluation should build the capacity of grantees to better design, manage, and evaluate their own efforts.

- Impact: Evaluation should advance the state of knowledge about how to best improve educational outcomes for students through educational leadership initiatives.

These principals should guide foundation staff as they identify the scope of the evaluation of any single program and as they develop and maintain collegial working relationships with grantee organizations.

The Purposes of Evaluation

The evaluation of a grant program may be undertaken for a variety of purposes. An evaluation may simply monitor a grant program's spending, or it may attempt to provide evidence that a leadership initiative is influencing student outcomes. As the W. K. Kellogg Foundation claims, evaluation can lead the foundation and those it supports to greater learning opportunities and more effective programs. Also, it can help the foundation set an example of thoughtful reflection on initiatives (W. K. Kellogg, 1998, i).

One important issue to consider when planning an evaluation is whether it is intended to be formative, summative, or both. A formative evaluation takes as its primary purpose helping the grantee to improve a program during the grant period; in this case, its main audience is the grantee organization. A summative evaluation assesses the overall impact of a program, either periodically or at its conclusion. The audience for summative evaluations may include a number of groups, ranging from the grantee and sponsoring foundation to other foundations, other potential sources of future funding, practitioners, researchers, and policymakers.

During interviews with Broad Foundation staff, RAND learned that the foundation wants its evaluation strategy to satisfy multiple purposes. Some of these purposes focus on formative

feedback to grantees, and some are more summative in the sense that they focus on the attainment of outcomes. The following list of important purposes is derived from discussions with Broad Foundation staff. Through its program evaluations, Broad wishes to:

- Monitor grant programs so that foundation staff have a rich understanding of the efforts of grantee organizations;

- Improve grant programs and build the capabilities of grantee organizations;

- Determine the best ways to implement successful leadership strategies so that their impact is sustained over time and so that they may be more easily adopted in other settings;

- Learn which programs work best so that the awarding of future grants will be well informed;

- Assess the entire portfolio of Broad Foundation grants to determine if the foundation supports the right mix of programs and to confirm that its grants promote foundation goals;

- Collect persuasive evidence of the efficacy of successful programs to encourage others to adopt them.

This list is intended to be comprehensive but not necessarily exhaustive. As time passes, the foundation may discover additional reasons to evaluate programs. Nonetheless, this list has guided the evaluation strategy that RAND has created. In addition, this list reflects RAND's intent to stress the strategic and intellectual usefulness of evaluation over its pragmatic reporting function, though both are important.

A strategy guided by these purposes maximizes the advantages of evaluation. More tangibly, the evaluation system should enable the foundation to be strategic in its giving. And in the long run, of course, evaluation plays a supporting role in helping the foundation achieve its over-arching goal of bringing about positive national changes in K-12 education.

Evaluation Strategy

The proposed Broad Foundation evaluation strategy consists of three components: a method for creating a goal-oriented description of a program, a point system for determining a program's appropriate level of evaluation, and a hierarchical evaluation system for collecting information

about the success of Broad grants. Using these tools, foundation staff can collaborate with applicants and grantees to identify the specific evaluation expectations for each program.

The strategy specifies that first, during the proposal stage, foundation staff work with an applicant to develop a diagram of a program's assumptions, activities, and goals as well as the links between these elements. The diagram, called a "logic model," usefully encapsulates a grant program in a manner that facilitates testing a program's clarity of focus and sequencing of activities. It also reveals possible methods for evaluating whether program activities are designed and executed in ways that can support the stated goals.

Next, foundation staff will apply what they learn from an applicant's logic model, along with additional grant information, to a point-based scoring equation to determine the scope of evaluation that is appropriate for the grant. The equation includes ratings for the size of the grant, the eventual use of evaluation findings, and the importance of the grant program in terms of the foundation's mission and overall strategy for improving education. In effect, the foundation rates where the grant fits in its portfolio and uses this rating to determine the scope of evaluation. For instance, staff will distinguish between grants that are high-profile, ambitious, multi-site efforts and those that are modest, one-time efforts to address problems in a single school district.

Once the grant's "score" is determined, its appropriate level of evaluation can be identified by referencing the hierarchical evaluation system presented below. This system encompasses five distinct levels of evaluation that the Broad Foundation may require of its various grantees. The levels range from a simple audit report on program resource use to a complex, scientifically reliable study to learn whether a program has led to systemic change in a region's or the nation's education system.

The following three sections describe each of these components in detail and suggest how they should be used by foundation staff.

Constructing Logic Models to Clarify Program Goals

A program logic model provides a diagram, or "road map," of how a program works—how the activities of a grant are assumed to ultimately affect the outcomes of interest. It "links outcomes (short- and long-term) with program activities/processes and the theoretical assumptions/ principles of the program" (W. K. Kellogg, 1998, 35). The logic model technique is particularly well suited for planning evaluations of programs where the outcomes of interest are not a direct result of the interventions and/or when these outcomes are difficult to measure directly. Both circumstances are common among leadership development initiatives.

Program logic models or similar models are used by a number of grant makers—including half of the foundations surveyed by Patrizi and McMullan (1998)—probably because of their simplicity. They are applicable to most activities and do not require specialized expertise, just thoughtful consideration of the proposal at hand. For Broad's purposes, the logic model is flexible enough to be used for programs of different scope, goals and approach, so the same tool can be used for projects that will require different levels of evaluation.

Logic models may be organized according to a variety of different schemes. The W. K. Kellogg Foundation identifies three major types: activities, outcomes, and theoretical logic models (W. K. Kellogg, 1998, 36-42). Perhaps the easiest of these to construct is an activities logic model, in which a program's main tasks are diagrammed chronologically. See Figure 1 for an example of an activities logic model of a hypothetical program offering math training to principals. Additional examples of logic models for other types of leadership development programs are provided in Appendix C.

Program description: A local university's principal academy sponsors Really New Math Training to help principals function as instructional leaders in mathematics.
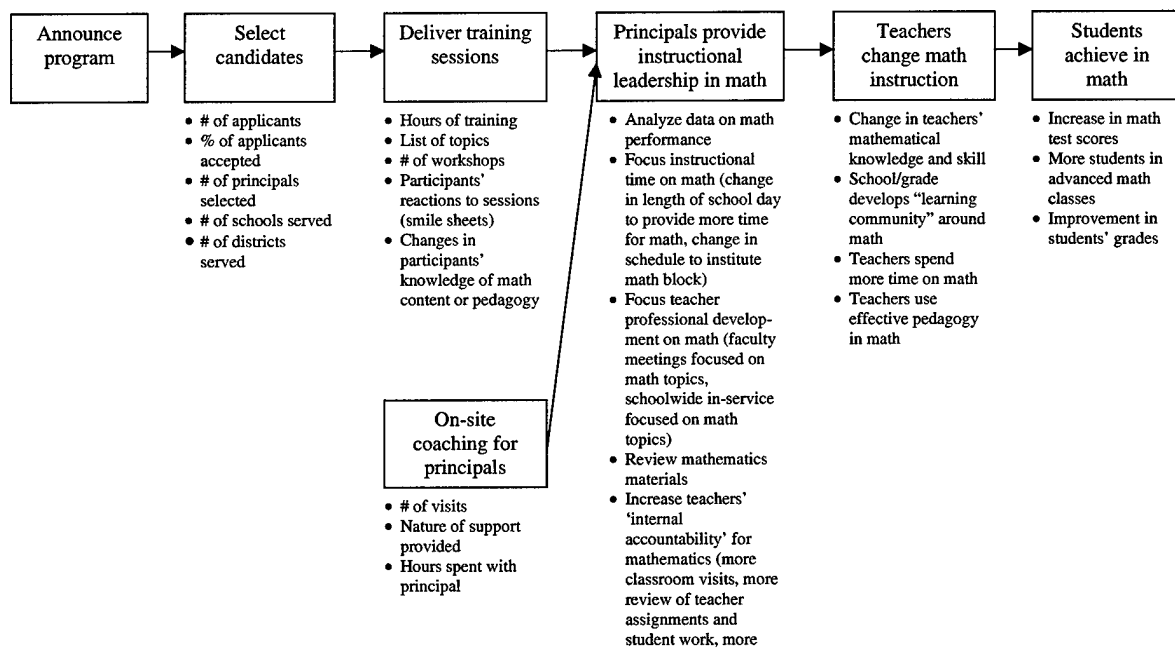

Figure 1. Logic Model for Really New Math Training Program

In the model, the boxes represent the major tasks of the program, and the arrows indicate relationship and sequence. The program would begin with a public announcement. Next, candidates for the training sessions would be selected, the training would be delivered to selected

candidates, and those who complete the training would receive on-site coaching in their schools. Once the training phase is completed, principals would become their schools' instructional leaders in mathematics. For example, they might institute changes in current math instruction, ask teachers to attend professional development sessions in mathematics, and review classroom materials on mathematics. As a consequence, teachers hopefully would change how they teach mathematics and, ideally, student achievement in mathematics would improve.

The bullet points below the activity boxes in the diagram list program indicators available for evaluation purposes at each stage of the program. Figure 1 diagrams a relatively simple program with clear relationships among tasks and one primary outcome; some programs may require more complex logic models with multiple lines of progression and multiple final outcomes. In those cases, the diagrams will be more elaborate (see Appendix C).

We propose that a logic model serve as the starting point for establishing evaluation expectations for each Broad Foundation proposal and grant. During the proposal development stage, applicants will develop draft logic models that describe proposed activities and how they will lead to desired outcomes. It is especially important that each applicant's model be as comprehensive as possible, listing all major activities and outcomes. Final outcomes such as improving student performance statewide should be included, even if they are ambitious and even if the program may have difficulty claiming sole or significant responsibility if the outcome is achieved. During the proposal stage, the purpose of the model is to identify and display the scope and breadth of the program as fully as possible, so outcome goals should not be limited by an awareness of their likelihood of achievement or long-term timeframe. Applicants will also identify indicators of success or evidence that would demonstrate completion and success for project tasks as well as outcomes. In Figure 1 these indicators are listed as bullet points below each activity. Next, a foundation staff member will discuss the logic model with the applicant, and together they will refine the diagram, adding in overlooked steps or indicators, clarifying terminology, and giving due consideration to the feasibility of program activities leading to desired outcomes.

Once a logic model is developed, it will be the basis for discussion between Broad Foundation staff and applicants concerning the appropriate level and type of evaluation. Staff and applicants can clarify their expectations about how they believe the proposed grant activities will make a difference in improving educational leadership and student achievement. When a proposal is funded and the evaluation plan is determined, the foundation and program officers will return to the logic model and discuss which indicators should be collected and used for evaluation purposes. The evaluation's level of intensity will be determined by a combination of factors, so it

is possible that some indicators initially listed on the logic model will be not be required when the evaluation is actually performed. We will return to these issues in the discussion of the hierarchical evaluation system beginning on page 13.

Throughout the life of a grant, logic models can be particularly useful for diagnosing problems and obstacles that may impede a program's success. They can help staff improve the program when early indicators suggest that the activities may not have had the desired impact. Once the program has reached a milestone or concluded, a logic model can help staff understand or explain why a program may have fallen short of its goals. For example, they can examine whether particular activities or processes considered to be crucial to produce intended outcomes actually occurred. If these events did not happen, then staff have an apparent explanation for the lack of effects, and can set about ensuring that the activity occurs in the future. If the activities or processes did occur, then staff can assess the quality of these components and reexamine their relevance to goals. By serving as an analytic tool, the logic model may help reveal which program activities need to be redesigned, substituted or abandoned. In this way, logic models may help grantees and foundation staff learn from mistakes and revise program plans, rather than being forced to discard a whole program.

Using a Point System to Determine Evaluation Needs

RAND's survey of foundations (see Appendix A) and those completed by Patrizi and McMullan suggest that grant-making organizations collect a lot of information from grantees to ensure accountability but seldom use the information. Of 21 foundations surveyed by Patrizi and McMullan, 12 report that almost all major initiatives and grants are evaluated; only eight of 21 report that at least half of grants, including small grants, are evaluated in some way.

Many foundations rely heavily on grantee self-evaluation and reports for information about grant activities. However, foundations often do not find the information useful because the quality of self-reported data is not always good, evaluations do not always answer questions of interest, results are often delayed, and the results supply just one data point in assessing a grant. One approach to improving the usefulness of evaluations would be to require more rigorous and timely evaluations as well as external evaluators. This approach, however, requires significant resources both in terms of time and dollars. In addition, rigorous evaluation is not appropriate for all grants. For example, some grants do not support large enough interventions to expect significant or measurable results.

RAND designed a point system as a guide for foundation staff to use when determining the level of evaluation for each grant. The system is intended to ensure that evaluations provide the most

useful information with the least cost and intrusiveness. It synthesizes three interdependent ratings about the appropriate level of investigation required given a grant's size, the eventual uses of evaluation findings, and the program's strategic importance to the Broad Foundation portfolio. As a proposal is working its way through the review and funding process, staff will score each grant on these three criteria. The ultimate "score" places a grant into one of five categories. These categories correspond to levels of evaluation.

RAND suggests that recipients of all but the most "trivial-sized" grants undertake the lowest level of evaluation by providing financial reports and statements. Many grantees should also undertake the next level, which documents completion of major activities. At the third level, a subset of grants will participate in an evaluation that examines the project closely enough for evaluators to identify the conditions and components necessary for program replication. A still smaller group of grantees will participate in evaluations that compare program outcomes across grants. Only a few grants—those that are large and of key strategic importance to the foundation's portfolio—will warrant evaluations that produce scientifically rigorous results. This level of evaluation will be reserved for programs for which Broad wants to use evaluation findings as evidence to demonstrate a program's success for policymakers and other critical audiences.

The point system rates each grant proposal according to following three characteristics:

1. **Size of Grant**: How large is the grant? Larger grants will warrant more intense evaluation. See Table 1.

Table 1. Size of Grant

Amount	Number of Points
Less than $50,000	1 point
$50,000 – $250,000	2 points
$250,000 – $1million	3 points
$1 million – $2 million	4 points
$2 million and up	5 points

2. **Uses of Evaluation Findings**: How will the findings of program evaluations be used? Will evaluation results be disseminated to the public via news stories and press releases to familiarize it with Broad's work? In such cases, anecdotal accounts of a program may be sufficient. Or might the grantee and foundation wish to use evaluation findings to "spread the word" about successful programs in more critical forums such as among

educators, researchers, foundations and policymakers? If so, the foundation may need scientifically reliable findings that will merit serious attention and bear scrutiny. One factor to consider in rating grants on this criterion is the project's potential "scalability," i.e., whether staff are hoping to apply the program or some of its activities in other districts or states. Scalability requires evidence about factors that facilitated success and an understanding of which components were most successful. Another factor of concern is the program's relevance to current policy issues. Strong evidence of results is required to convince policymakers, so foundation staff should award the maximum number of points when the goal is to affect policy.

See Table 2 to translate the intended uses of evaluation findings into points to help determine the intensity of a program's evaluation. The four items in this table stem from RAND's interviews with foundation staff, and they are intended to be understood as cumulative. In other words, any evaluation whose findings might be use to implement a program in additional locales could also be used to inform the public.

Table 2. Uses of Evaluation Findings

Uses of Evaluation Findings	Number of Points
Inform public	1 point
"Travel" to other schools/districts	2 points
Provide basis for securing funding from other sources	3 points
Provide evidence to support policy change	4 points

3. **Strategic Importance of Grant to Foundation**: To what extent does a grant address key strategic goals of the Broad Foundation? Is it part of a multi-site, multi-grant initiative? Program officers should award grant programs a score out of five possible points based on their degree of importance to the work and to the reputation of the foundation. See Table 3.

Once a grant is rated on these three criteria, staff tally the points. The total points awarded for each grant proposal suggest a corresponding level of evaluation. See Table 4 for a cumulative description of the point system. Evaluation levels are described in the following section.

Table 3. Importance of Grant

Program's Level of Importance	Number of Points
Relatively unimportant (e.g., a small or short-term program)	1 point
Marginally important (e.g., an experimental program)	2 points
Of average importance	3 points
Relatively important	4 points
Vital to Broad's public reputation (e.g., a high-profile or large, multi-year program)	5 points

Table 4. Determining the Right Level of Evaluation: Score Sheet for Grants

Program name: _____

Program dates: _____

Grantee organization: _____

Criteria *(Assign points once in each category)*	Number of Points
Size of Grant	
Less than $50,000	1 point
$50,000 – $250,000	2 points
$250,000 – $1 Million	3 points
$1 Million – $2 Million	4 points
$2 Million and up	5 points
Uses of Evaluation Findings	
Inform public	1 point
"Travel" to other schools or districts	2 points
Provide basis for securing funding from other sources	3 points
Provide evidence for policy change	4 points
Program's Level of Importance	
Relatively unimportant (e.g., a small or short-term program)	1 point
Marginally important (e.g., an experimental program)	2 points
Of average importance	3 points
Relatively important	4 points
Vital to Broad's public reputation (e.g., high-profile) large, multi-year program	5 points
Total	
Recommended Level of Evaluation	*Total Points*
Resource Use	3 – 5 points
Activities	6 – 8 points
Outcomes	9 – 14 points
Shared Outcomes	at Broad's discretion
Systemic Change	at Broad's discretion

Matching the Level of Evaluation to the Grant

RAND conceived of the universe of Broad grant evaluations as fitting into five levels of intensity, ranging from a simple accounting of grant resource use to elaborate, scientifically valid evaluations that might reveal changes to education systems. These levels can be usefully pictured as horizontal slices of a triangle (see Figure 2).
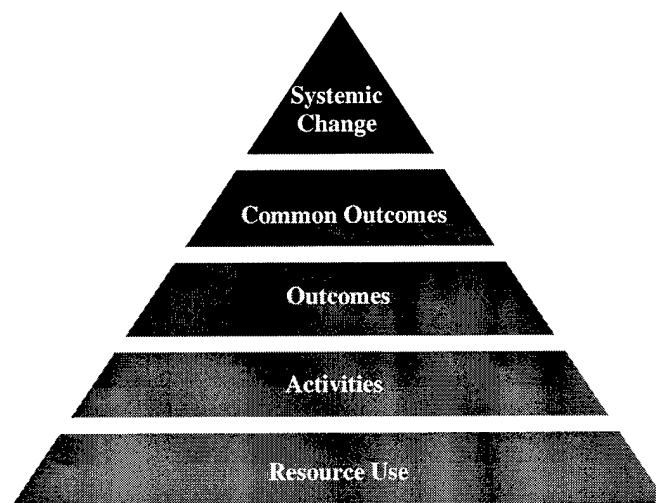
Figure 2. Evaluation Strategy for Broad Foundation Grants

The triangle illustrates the core principles of the foundation's evaluation approach. Its levels reflect the foundation's commitment to appropriate evaluation. The triangle's tapering sides suggest that fewer and fewer grantees or programs will be evaluated in the higher levels. The base level—an audit of resource use—is necessary for accountability and is expected of almost all grants, while the capstone of the pyramid—an evaluation of systemic change in the environment—is necessary for gauging the larger impact of a grant program on education policy and achievement and is expected of few grants. This hierarchy inherently suggests that moving from the base of the triangle toward the tip, level by level, signals heightened evaluation activity with each stage; grant programs evaluated at the higher levels will need more resources for their more intensive evaluation requirements. The lower three levels are project-specific, while the top two address issues that transcend individual projects. In addition, the levels are cumulative. For example, a level-3 outcomes evaluation would also include the evaluation activities performed for levels 1 and 2.

A description of the five evaluation levels follows.

Level 1: Resource Use

Description of level: The purpose of level 1 is to "audit" grants and ensure basic financial accountability. A level-1 evaluation provides the Broad Foundation with rudimentary information about grantees' use of funds. While not detailed enough to scrutinize grantees' every spending decision, the financial reports can raise concerns if spending appears to be off-track or inappropriate.

Evaluation activity: Grantees will provide semi-annual financial reports of the grants which include the grant budget, grant expenditures to date, and grant expenditures in the last period. For grants that require matching funds, grantees must also submit semi-annual reports of the amounts and sources of matching funds. Grantees will also submit their organization's balance sheets annually, as reported in the organization's annual report. Grant-making organizations routinely require such financial reports.

Responsibility for reporting: Grantees will prepare and submit financial reports and statements.

Level 2: Activities

Description of level: The purpose of level 2 is for the foundation to learn about the activities undertaken by its grantees. A level-2 evaluation provides the Broad Foundation with information to monitor whether programs are implementing proposed activities. The information can also be used to help both foundation staff and grantees assess whether the program is following its logic model and if program activities are occurring as planned. In large part, the reflection and discussion prompted by a level-2 evaluation will help the grantees strengthen ongoing programs, playing a formative role. Finally, staff may use grantee-provided activities narratives to describe foundation-funded activities to the board and the public. This level of information can generate "doing good" lists, which, in some cases, may provide adequate evidence of the success and value of a program.

Evaluation activity: Semi-annually, grantees will provide a short narrative (under 20 pages) describing program activities conducted during the last six months. The activity reports may also include process measures of the activities, such as services delivered and the number of people served. Since these measures are the direct result of project activities, the data can be collected fairly easily by the grantees, often while an activity such as a workshop or seminar is taking place. The specific measures for inclusion will be negotiated between foundation staff and grantees during proposal development. Program logic models will be useful in identifying measures; for example, the bullet points under the activity boxes in Figure 1 are a starting point

for developing appropriate evaluation measures for this level. When possible, the foundation will ask grantees to use measures common to multiple grant programs so that it can aggregate the figures to report the scope of the foundation's impact.

Many grant-making organizations require level-2 reporting routinely; Patrizi and McMullan found that at least half of foundations required this type of reporting, usually done by the grantee (Patrizi and McMullan, 1998). Most grant-making organizations develop indicators on a case-by-case basis in order to tailor the measures to the specific activities of the grantees. However, unique measures make aggregation and comparison of the indicators difficult; for this reason, we encourage the foundation to ask grantees to use common measures whenever possible.

Responsibility for reporting: Grantees will prepare and submit narratives. Grantees and foundation staff will collaborate to determine the process and output indicators to be reported.

Level 3: Outcomes

Description of level: The purpose of a level-3 evaluation is to monitor both the progress and the outcomes of grants as a basis for a decision about continued investment. Grantees may also use the information for continuous improvement of their programs and to make a case for additional funding.

Evaluation activities: Level 3 includes a range of activities from tracking implementation of the program to collecting and reporting data about outcomes. Sources of data for this level of evaluation may include surveys, focus groups, interviews, databases and documents. The type of data collection and analysis, as well as the timing of the outcomes evaluation, will need to be jointly determined by foundation staff and the grantee. These activities may be carried out by the grantees; in some cases, however, the Broad Foundation may choose to hire an external evaluator to conduct the level-3 evaluation in order to increase the credibility of the evaluation results, to minimize the burden on the grantee, and to provide capacity necessary to carry out a rigorous and technically challenging study.

Responsibility for reporting: Foundation staff will work with the grantee to determine the type of information that will be collected. Once again, the program's logic model may provide a heuristic for determining who should be surveyed and what information or documents might be collected. Some grantees may be well equipped to handle these complex tasks, while others may require external assistance. The Broad Foundation may retain an external evaluator in order to provide more objective feedback about the strengths and weaknesses of the program.

Regardless of the timing of the specific level-3 outcomes evaluation, grantees in this category will be required to submit spending and activities reports every six months, as described under levels 1 and 2.

Level 4: Common Outcomes

Description of level: The purpose of a level-4 evaluation is to compare outcomes of similar projects. Projects may be similar because the same initiative is being implemented in multiple sites, or, alternatively, different programs may use the same mechanisms for change or seek similar outcomes.

The foundation may have an interest in pursuing a level-4 evaluation to learn which site(s) of a multi-site and multi-grantee effort are successfully implementing a program. Such an evaluation may also reveal key information about what elements lead to program success or which obstacles are most likely to undermine it. Insights gained from a level-4 evaluation may help the foundation to plan its giving strategy for the future and to advise applicants about successful strategies as they are developing logic models and proposals for new programs. In addition, the foundation may use a level-4 evaluation to consider which program among a group of different interventions with similar goals is the most beneficial.

Participating grantee organizations should also benefit directly from a level-4 evaluation. The effort should provide grantees a broader perspective on the merit of their individual programs and spur creative contemplation about the best ways to achieve shared program goals. Level-4 evaluation should point to which activities within programs appear to have the broadest, deepest, or quickest impact. More specifically, grantees may learn about more effective approaches to implementation from each other through a level-4 evaluation. For example, grantees may gain insights into how to improve the sequence of activities, their timing, or the group size for an event from learning about how other organizations are structuring their interventions. In other words, the participants should *not* experience a level-4 evaluation as a ranking process; rather, it should be a learning opportunity—and perhaps even an honor—to be included in such an undertaking.

RAND's conception of level-4 evaluations shares some of the purposes and features of the "cluster evaluations" conduced by W. K. Kellogg, The Pew Charitable Trusts, and others. In large part, these consist of assessments of a foundation's portfolio of grants (Patrizi and McMullan, 1998). For example, W. K. Kellogg assesses similarly grouped projects "to determine

how well the collection of projects fulfills the objective of systemic change" (W. K. Kellogg, 1998, 17).[1] W. K. Kellogg carefully defines this particular approach to evaluation:

> Cluster evaluation focuses on progress made toward achieving the broad goals of a programming initiative. In short, cluster evaluation looks across a group of projects to identify common threads and themes that, having cross-confirmation, take on greater significance. Cluster evaluators provide feedback on commonalities in program design as well as innovative methodologies used by the projects during the life of the initiative (W. K. Kellogg, 1998, 17).

The Pew Charitable Trusts believes the advantage of cluster reviews is that they provide "strategic information on a coherent set of grants that represent a considerable investment" (W. K. Kellogg, 1998, 19).

Evaluation activities: Level-4 evaluations will require that either the grantee or external evaluators collect data on common indicators of program processes and outcomes. The complexity of the data and concerns about comparability of data elements across grantees will determine whether the indicator data is collected by grantees or external evaluators. Once collected, foundation staff or external evaluators will draw comparisons among the various sites or programs.

A productive approach to organizing a level-4 evaluation is to use or develop a set of common metrics for all grantee organizations that participate. These metrics would reflect the shared outcomes the projects sought to achieve as well as common procedures that were salient to all models. To a degree, it is helpful to conceive of these metrics as deriving from a common "meta-logic model," that is, one that can be applied to all the relevant grant projects with nearly equal ease and without distorting their activities and goals.

Interestingly, grant-making organizations rarely use common metrics to assess their grantees. Rather, most organizations and grantees develop indicators and performance measures on a case-by-case basis. In fact, a Community Wealth Ventures study for the Morino Institute found that "few [grant-making] organizations possess specific measures to calculate investment returns. Most fund executives noted that return measures are still being determined and will probably be formulated on a case-by-case basis" irrespective of the organization's giving strategies (e.g.,

---

[1] W. K. Kellogg's cluster evaluations include features of both Level-4 and -5 evaluations in RAND's evaluation strategy. We discuss them here because of their emphasis on evaluating groups of projects.

whether it is a venture philanthropy or a traditional foundation) (Community Wealth Ventures, 2000, 39). Perhaps common metrics are rare because they demand a higher level of planning and coordination than that typically found. In addition, grantees often want to portray their programs as unique. Nonetheless, we believe that in some cases it may be worthwhile for the foundation to invest the staff time and resources necessary to develop common metrics for a family of grants.

At the same time, comparing projects using common metrics must be undertaken with caution since programs that are not simply replications are likely to have different approaches and address different aspects of a problem. The task of identifying common measures may be more complex in practice than it appears in theory. In some cases, it may be necessary to find alternative ways of measuring the same general feature. In addition, foundation staff should be careful about interpreting comparisons based on the common metrics. The variables measured may have different salience for different projects. For example, comparing changes in student achievement test results may be appropriate for some groups of grants but not for others; the range of activities for the grants included in a level-4 evaluation would determine which outcomes might be compared.

The use of common metrics will require careful evaluation conceptualization and planning on the part of foundation staff or external evaluators. Unfortunately, the literature does not identify a concise set of common metrics that can be used to measure the issues and reforms of interest to the Broad Foundation. A review of the literature on educational governance, management, and labor relations yielded only one source that might be used to develop a set of indicators—for grants promoting "new unionism" (Kerchner and Koppich, 1998). Given this dearth of ready models to apply or refine, staff and others may find it helpful to refer to studies that have developed common indicators to assess a range of grant programs. One such evaluation is described in (Bodilly et al., 1998). In this study, researchers designed a strategy for evaluating six distinctive whole-school reform designs and their implementation.

The way programs are classified and selected for a level-4 evaluation may bear thought. Which grouping strategy will be most amenable to evaluation, and which will provide the best insights? For example, programs might be grouped according to their leadership focus (i.e., governance, management, or unions). Alternatively, they could be grouped according to the goals of program activities (i.e., recruiting leaders, redefining roles, improving leadership capacity, creating incentives, and honoring success), or by a more limited combination approach (e.g., programs that improve the leadership capacity of school board members). When making this decision, staff should also consider whether the foundation is likely to evaluate this group of programs at level 5

as well because measuring or assessing "impact" may be more telling for particular groups of programs than others.

Responsibility for reporting: We expect that in some cases, foundation staff will leave the options for level-4 and -5 evaluations open, perhaps even for a period of several years, and then decide to invest in them once their value becomes apparent. Contingency planning for level-4 and -5 evaluations might, then, prove useful and would amount to including additional questions in lower-level assessments. Although individual grantees will not bear responsibility for all data collection in a level-4 evaluation, they may be asked to collect some data suitable for use in a level-4 evaluation. Much of this information may already be available in the lower-level reports (levels 1-3). It is likely, however, that foundation staff and/or external evaluators will conduct additional data collection, analysis, and reporting for level-4 evaluations because of their complexity and sensitivity.

Level 5: Systemic Change

Description of Level: The purpose of level 5 is to understand the systemic impact of Broad Foundation initiatives. It assesses systemic effects, not individual projects and not just variation among projects. The emphasis on impact, an explicitly summative concept, distinguishes level-5 from level-4 evaluations, which may have a range of formative and summative purposes. "Systemic change" refers to the foundation's affect on the education system. Of course, that concept can be taken to have a range of meanings, including impact on the education system of a district, region, or state; its influence on educators and education researchers and policymakers at the state and national levels; and, perhaps, its impact on the ability of schools to achieve the public goals of education such as assuring that students master basic skills, become lifelong learners (e.g., learning how to learn), and understand the nation's government and values well enough to be good citizens. Consequently, it may be best for the foundation to refine its definition of "systemic change" each time it conducts a level-5 evaluation.

Other basic choices foundation staff need to make early in the process of planning a level-5 evaluation are the kind of claims that it would like to make based on the findings and whether these findings will be used internally or externally. If staff are conducting a level-5 evaluation as a means of foundation self-evaluation for in-house purposes, including some level-5 indicators and questions in a level-4 evaluation plan may suffice. Also, whether or not a family of grant programs has a broad or significant impact may sometimes be easy to determine. For example, it may be relatively simple to track a grant initiative's media coverage or its discussion in public debate or legislation.

If, however, staff desire evidence to make public claims about the efficacy or value of its programs and grant strategies, then a scientifically rigorous, external evaluation will probably be required. Similarly, determining whether programs that have been funded, adopted in multiple sites, and implemented over a period of years have actually changed the traditional (and presumably less-effective) practices of school leaders and those who hire, educate, and supervise them, will merit empirical investigation. In any case, a level-5 evaluation will have real pay-off to foundation staff by providing them with concrete data about how one set of efforts fits into the foundation's master plan; findings can be folded into a business plan and direct future grant-making priorities.

Evaluation activities: Level-5 evaluations build on level-4 evaluations (either previous or concurrent), hence the design of the level-4 evaluation may prove critical to the focus of the capstone level-5 effort. (On occasion, a level-3 evaluation may lead to a level-5 evaluation directly, but we assume this will be a rare event.) Consequently, the classification of programs in a level-4 evaluation may direct or constrain the questions and findings of a related level-5 evaluation. Once again, we recommend contingency planning for level-5 at earlier stages because it can save the foundation and its grantees time, effort, and resources. However, as in other cases, data collected for a level 1-4 evaluation may inform the level-5 evaluation.

Key questions that a level-5 evaluation should address include:

- **What impact has the program had on the education system?** Have our programs leveraged other resources for our grantees? Have programs been replicated? Have we seen a shift in the culture of educational institutions as a result? Is there a change in public support for education?

- **Who else is addressing this issue?** Are our approaches similar or different? Is there anything we can learn from others' approaches to this issue?

- **Has the program precipitated interest from others?** For example, have foundations, government agencies, legislators, or executives become aware of a program and expressed interest in more widespread implementation?

- **Has the program changed how educators, researchers, policymakers, and/or the public think about an issue?** Have state legislatures or Congress drawn on a program's theory, activities, or outcomes as they debated issues or crafted legislation? Has the program been endorsed by any individuals or organizations? Are parents and the media aware of the program?

Responsibility for reporting: Level-5 evaluations conducted purely as in-house exercises in self-reflection may be carried out by foundation staff. However, it may also be useful to have external evaluators perform a more formal review every five or 10 years. Several foundations—Pew, The Heinz Endowments, W. K. Kellogg, Rockefeller Foundation, Eli Lilly and Company Foundation, Charles Stewart Mott Foundation, and the McKnight Foundation—have used similar approaches for strategic planning, and they use external evaluators for these efforts. Evaluations that aspire to assemble evidence for public dissemination about the foundation should be conducted by external evaluators to assure credibility.

Using Evaluation Information

Reviewing and discussing interim or completed evaluation reports internally and with grantees should be a priority for Broad Foundation staff. Staff should assess the value of the information that they have collected from grantees and determine if they are indeed using the data. The purpose of dedicating foundation and grantee resources to evaluation is to monitor grants and inform foundation decisions about continuing investments, promoting successful programs in the media, supporting program replication, advocating for policy changes and reviewing foundation giving strategies. If, over a period of time, staff discover that they are not making use of evaluation findings, they may want to either scale back grantee evaluations or develop explicit dissemination plans for those findings through internal review or through external reports, newsletters, and web pages.

References

Bodilly, Susan J. et al. (1998). *Lessons from New American Schools' Scale-Up Phase: Prospects for Bringing Designs to Multiple Schools*. Santa Monica, CA: RAND.

Broad Foundation (a). "About the Foundation" in the Broad Foundation web site. Los Angeles, CA: Broad Foundation. http://www.broadfoundation.org/about.htm. Accessed 30 April 2001.

_____. (updated brochure, b) *The Broad Foundation: Raising Performance, Supporting Innovation, Making Success Visible, Developing New Leadership*. Los Angeles, CA: Broad Foundation.

Brookings Institution Governmental Studies Program and Harvard University Project on Effective Interventions (1998). *Learning What Works: Evaluating Complex Social Interventions, Report on the Symposium Held on October 22, 1997*. Washington, DC: The Brookings Institution.

Cohen, David and Heather Hill (February 1998). *Instructional Policy and Classroom Performance: The Mathematics Reform in California* (No. RR-039). Philadelphia, PA: University of Pennsylvania Consortium for Policy Research in Education.

Coleman, James (1996). *Equality of Educational Opportunity: Summary Report*. Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education.

Community Wealth Ventures, Inc. (2000). *Venture Philanthropy Landscape and Expectations*. Reston, VA: Morino Institute Youth Social Ventures.

Emerson, Jed, Jay Wachowicz, and Suzi Chun (2000). "Chapter 8: Social Return on Investment: Exploring Aspects of Value Creation in the Non Profit Sector," in Roberts Enterprise Development Fund (Ed.), *Social Purpose Enterprises and Venture Philanthropy in the New Millennium*, Volume 2. San Francisco, CA: Roberts Enterprise Development Fund. 132-173.

Ferguson, Ronald (Summer 1991). "Paying for Public Education: New Evidence on How and Why Money Matters." *Harvard Journal on Legislation*, Vol. 28, No. 2.

Gates, Susan, Karen Ross, and Dominic Brewer (2000). "School Leadership in the Twenty-First Century: Why and How Is It Important?" *Leading to Reform: Educational Leadership for*

the 21st Century. School Development Outreach Project. Oak Brook, IL: NCREL, North Central Regional Educational Laboratory.

Karoly, Lynn A. et al. (1998). *Investing in Our Children: What We Know and Don't Know About the Costs and Benefits of Early Childhood Interventions*. Santa Monica, CA: RAND, MR-898-TCWF.

Kaplan, Robert and David P. Norton (September-October 1993). "Putting the Balanced Scorecard to Work." *Harvard Business Review*, 134-147.

Kaplan, Robert and David P. Norton (1996). *The Balanced Scorecard*. Boston, MA : Harvard Business School Press.

Katzir, Dan (July 25, 2000). "Opening Remarks for the Panel on Foundation Strategies for Strengthening Education Leadership." Council of Chief State School Officers Conference, Wilmington, DE.

Kerchner, Charles and Julia Koppich (1998). *A Union of Professionals*. http://63.197.216.234/crcl/mindworkers/journart.cfm?pid=6. Accessed 16 August 2000.

Patrizi, Patricia and Bernard McMullan (December 1998). *Evaluation in Foundations: The Unrealized Potential*. Battle Creek, MI: W. K. Kellogg Foundation. http://www.wkkf.org/documents/wkkf/evalpotential/evalpoten1.asp

Rimel, Rebecca W. (undated). "President's Message: Accountability" The Pew Charitable Trusts, Philadelphia, PA, http://www.pewtrusts.com/about/about.cfm. Accessed 5 December 2000.

Roberts Enterprise Development Fund (undated). *SORI Reports: Overview and Guide*. San Francisco, CA: The Roberts Enterprise Development Fund. http://redf.org/download/sroi/red.overview.pdf. Accessed 30 April 2001.

Twersky, Fay (2000). "Webtrack and Beyond: Documenting the Impact of Social Purpose Enterprises: Profile of a Management Information System, in Roberts Enterprise Development Fund (Ed.), *Social Purpose Enterprises and Venture Philanthropy in the New Millennium*, Volume 2. San Francisco, CA: Roberts Enterprise Development Fund. 119-123.

W. K. Kellogg Foundation (1998). *W. K. Kellogg Foundation Evaluation Handbook.* Battle Creek, MI: W. K. Kellogg Foundation.

Word, Elizabeth et al. (1990). *The State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project, Final Summary Report, 1985-1990.* Nashville, TN: State Department of Education.

Appendix A:

Description of the Evaluation Strategies of Selected Foundations

RAND conducted an informal survey of other philanthropic organizations' evaluation efforts. They were chosen on a variety of factors. Broad Foundation staff suggested some of the organizations (e.g., the Panasonic Foundation); others were identified through RAND's literature review of how philanthropic organizations evaluate the grant programs they support.

Twelve organizations were selected based on their reputations for being innovative and on other characteristics they share with the Broad Foundation. This sample includes both large and small foundations (e.g., Echoing Green Foundation and Edna McConnell Clark Foundation) as well as five foundations that operate with a traditional approach and seven that view themselves as venture philanthropists, or organizations that invest in social entrepreneurs. The selected organizations share the following characteristics with the Broad Foundation:

- An emphasis on education;

- A reputation for being "results-oriented;"

- A strategy that invests in leadership development;

- A reputation for assessing the impact of its funding.

Table A.1 lists the foundations RAND contacted, the amount of their endowment or the venture capital they have to invest, their interests, their approach to grant making, and the title of the foundation representatives who were interviewed.[2]

We conducted unstructured interviews with representatives from each foundation and supplemented this information with documents from the foundations and from their web sites. Our object was to talk with the staff member who was most knowledgeable about the foundation's evaluation goals and process. Respondents included directors of evaluation, program managers, and foundation officers. The following topics were discussed:

- Purpose and use of evaluation information

- Evaluation measures

---

[2] Respondents were assured anonymity.

- Evaluation process

- Capacity and support for evaluation

This appendix summarizes our impressions of the foundations' approaches to evaluation based on all of the materials and information made available to us.

Table A.1. Foundations Interviewed

Foundation	Size of Endowment, Assets, or Grants	Focus(es)	Approach	Respondent
Ashoka	Endowment: $7 M	Social problems	Venture philanthropy	Associate director in the venture department
Colorado Trust	Assets: $394 M	Health, families	Traditional	Evaluation expert
Echoing Green	Endowment: $1 M	Public service	Venture philanthropy	Director of assessment
Edna McConnell Clark	Grants: $27.7 M (1998)	Youth development	Venture philanthropy	Director of assessment
Entrepreneurs'	Liquidity: $11 M (2/2000)	Education, youth development	Venture philanthropy	Director of venture philanthropy
James Irvine	Grants:$75 M	Social, physical, economic life in California	Traditional	Associate director of evaluation
Panasonic	Endowment: $10 M	Education	Traditional	President
Pew Charitable Trusts	Assets: $4.9 B Giving: $250 M (1999)	Education	Traditional	Program officer
Roberts Enterprise Development Fund		Low-income workers	Venture philanthropy	Managing director
Robin Hood	Grants: $12 M (1999)	Health, schools, neighborhoods	Venture philanthropy	Managing director
Social Venture Partners	Grants: $1 M (2000)	Education	Venture philanthropy	Director
William and Flora Hewlett	Grants to education: $22 M (1998)	Education, the arts, the environment, communities	Traditional	Program officer for education

Foundations' approaches to evaluation generally depended on their view of grant making. In our discussions with foundation representatives, we found two categories of foundations, each with different goals and evaluation strategies: traditional foundations and venture philanthropy funds.

Traditional foundations identify philanthropic areas or topics in which they are interested such as education, childhood health, or alternative school designs. Some describe specific initiatives they seek to promote and others leave problem definition up to organizations submitting proposals. Typically, grantees are agencies (e.g., school districts) or institutions (e.g. universities) that submit proposals to study an issue of interest or to pilot a strategy to ameliorate a problem. Evaluation focuses on either the role of the foundation in implementing its agenda, or the process of implementing the funded program (e.g., reporting how many students participated in an after-school enrichment and reading program).

In the last decade, several other approaches to resolving public welfare issues have taken hold, including venture philanthropy and the funding of social entrepreneurs. Venture philanthropists support individuals or non-profit businesses or organizations that seek to improve the quality of life of the disadvantaged. Typically, these organizations attempt to provide jobs to individuals in disadvantaged communities. Following a venture capital model, grantees present a business plan and set milestones and performance measurements to be achieved. The foundation staff are actively involved with the organization in developing the plan and its implementation. The foundation's interest is in developing the capacity of the grantee to organize and operate a successful business, grow it, and sustain its operation. Venture philanthropy also supports individuals, often called "social entrepreneurs," in their efforts to form an organization or provide services to address social problems. One project, for example, involves establishing a type of community-run computer school in an urban slum. For this philanthropy, program evaluation focuses on the role of the foundation in assisting the development of the funded program/individual, and in the achievement of milestones and measures established in the business plan.

Common Themes

While individual foundations varied in their approach to evaluation, a number of common practices were apparent. We discuss these common approaches below and note the exceptions that may expand our understanding of evaluation strategies. Table A.2 summarizes the findings of this research.

Table A.2. Summary of Evaluation Purposes and Procedures

	Primary Purposes	Evaluation Measures Used	Evaluation Schedule	Evaluation Costs	Capacity and Support
Ashoka	• Learn about field • Keep in touch with grantees • Tailor assistance • Learn about program impact • Report to donors	• Narrative describing activities	• Semi-annual reports from fellows • Selected review of fellows five years after grant period ends		Connects fellows with other organizations or fellows who can be of assistance
Colorado Trust	• Learn about impact • Judge effectiveness in working with grantees • Learn about progress of an initiative • Help grantees enlist other funders	• Process • Output • Beginning to supply outcome measures • Developing database for sharing among multiple sites	• Semi-annual progress reports • Two site visits per year	Varies by initiative. Typically, 10-50 percent of grant goes to evaluation	Uses an evaluation firm to help grantees develop individualized sets of outcome indicators. Select outside evaluator by competitive RFP for each initiative
Echoing Green	• **Short-term**: Success reaching objectives • **Long term**: Social impact of grants • Organizational and leadership development of grantees	• Detailed budgets • 6-month logic models with objectives for which grantees are accountable • Business plans	• Planning meeting every six months • Survey past grantees		Provides consultants who help to organize evaluation measures
Edna McConnell Clark	• Foundation effectiveness • Institutional learning • Sharing information about youth development with others	• Business plan with milestones • Outcomes • Multi-purpose database	• Clark portfolio mangers report weekly		Portfolio managers are deeply involved in monitoring progress towards milestones. Also uses outside consultants

Entre-preneurs'	• Look at process and outcomes to scale up the organization • Build capacity of grantees • Inform funding community	• Performance measures • Milestones • Business plans	• Formal report every six months • Site visits every two weeks		Foundation staff and consultants help grantees with evaluation
James Irvine	• Inform foundation • Inform foundation and professional communities	• Process • Output • Measurable objectives, timelines, progress indicators	• Annual reports	Evaluations of multi-grant initiatives: $7-$10 million over three years	Two-member in-house evaluation unit assists program staff and offers technical assistance to grantees
Panasonic	• **Short term:** Learn about impact • Judge effectiveness in working with grantees • Learn about the progress of an initiative • **Long term:** Value added, foundation's role as partner	• Process	• Ongoing formative evaluation • Annual site review • Five-year formal assessment of school district plan		On-site consultant at each school district
Pew Charitable Trusts	• Feedback to program staff on key indicators • **Long term:** Impact and theory underlying a strategy in a cluster of programs	• Binding monitoring plan • Measurable objectives with process and output benchmarks	• Semi-annual financial reports • Annual narratives • Evaluate of portfolio of grants in single area every 5-7 years		Professional in-house evaluation team. Plans are developed by program staff. Hires external evaluators
Roberts Enterprise Develop-ment Fund	• **Short term:** Commun-ication to determine funding needs • **Long term:** Appraise success of the fund	• Establishing database to measure operational and social outcomes • Monthly financial indicators	• Meet monthly • Monthly financial reports		Social outcomes expert works with grantee to develop database and data collection plan. Also uses outside consultants

Robin Hood	• Help develop programs	• Outcomes such as attendance, reading and math scores, parent involvement	• Annual review	Paid by the foundation	Uses independent evaluators
Social Venture Partners	• Determine need for additional funds • Inform grantees so they can improve program	• Outcome Milestones: Program-matic and organiza-tional capacity building • Developing system to aggregate outcomes from individual grantees	• Annual review		Foundation works with grantees to develop goals, objectives, marketing plans, infrastructure, etc.
William and Flora Hewlett	• Identify assistance and funding needs	• Adding outcome measures for large grants • Narrative and financial reports	• Annual reports	Six percent of Bay Area School Reform Collaborative spent on evaluation	Uses outside evaluator for large initiatives. Foundation staff may help design study and evaluation, and assist smaller grantees in finding evaluators

Purpose and Use of Evaluation Information

Traditional foundations such as the Colorado Trust, Panasonic, and the James Irvine Foundation use evaluation to learn about their impact in specific areas, judge their effectiveness in working with grantees, and inform themselves about the progress of an initiative. For the most part, they rely on process and output measures rather than outcome measures. Usually, funding continuation is not based on evaluation findings. However, there seems to be a movement towards more outcome-based evaluation measures.

Interestingly, several foundations are finding that grantees want more rigorous evaluations. For example, grantees of the Colorado Trust want outcome evaluations both to assess their impact and also to provide information for future funders. In response, the Trust has changed its policy and will now provide technical assistance through an outside evaluation firm to help each grantee develop its own set of indicators. The William and Flora Hewlett Foundation is also changing its

evaluation strategy, particularly for large grants. It is devoting more resources to evaluation and taking a more active role in designing studies and working with evaluation teams.

Though they also view evaluation developmentally, venture philanthropists approach evaluation differently from traditional organizations. Venture philanthropists use evaluation primarily for determining how to assist an organization in developing the expertise to operate its business and assess its success as well as to determine the amount of additional financial assistance the organization will require from the foundation. Grantee organizations set and must meet specific milestones and outcome measures. Grantees are encouraged to obtain a broad stream of financial support, and to incorporate the measures of all funders in their evaluation strategy. For instance, one non-profit organization that the Roberts Enterprise Development Fund (REDF) contributes to has 170 other funders each with outcomes that they have specified and want to see achieved besides the outcomes REDF has specified.

RAND talked with two "venture capital" foundations that direct their support to individuals seeking to promote social change. Ashoka, a foundation supporting fellows worldwide, sees itself as a professional association of entrepreneurs. It uses an extensive and intensive pre-grant selection process to identify fellows. Once selected, fellows report semi-annually. One of these is an open-ended narrative describing the grantee's activities; the other consists of fellows' responses to a specific set of questions. The questions include the following:

- How many people are affected by the work of the fellow?

- What legislative changes have occurred?

- Has the innovation spread to other institutions?

- What services were provided to the fellow by Ashoka?

Together, the two reports are used to learn about the field, keep in touch with fellows, tailor assistance, and report activities to donors. Notably, the foundation tries to be as *non-evaluative* as possible, giving grantees leeway to frame issues and strategies in their own ways.

The Echoing Green Foundation invests in independent projects that are in the start-up phase. Grantees are required to develop detailed budgets and six-month logic models with specific objectives to which they are held accountable.

Evaluation Measures

Few of the foundations we interviewed use a common set of evaluation measures for all grantees. Typically, an evaluation is geared to the grantee's project and often designed in consultation with the grantee. Traditional foundations focus on process measures, while the venture philanthropists focus on outcome measures. However, according to a program officer at Pew Charitable Trusts, there is a movement among traditional foundations towards more rigorous outcome evaluations, especially with large grants which are "rich in learning opportunities."

Monitoring information is typically provided to the foundation in annual or semi-annual financial reports and narrative reports of activities. Some grantees are held accountable to a specific set of objectives. At Pew, for example, once a grant is approved, program staff develop a monitoring plan which is a binding document. The plan contains measurable objectives with benchmarks. The nature of the objectives and rigor of the benchmarks are developed by program staff and are based on the proposal and the funded work. At Entrepreneurs' Foundation, a venture philanthropy, the Board of Directors, staff, and grantee organization work together to develop performance measurements and milestones that will be incorporated into the grantee's business plan. Echoing Green's director of assessment and outside consultants provide one-to-one assistance to fellows to develop comprehensive logic models with evaluation measures. Fellows continue to develop six-month plans throughout their grants, and they are held accountable to them. The Robin Hood Foundation's school initiative in New York hires an independent evaluator to evaluate each recipient annually. The evaluation focuses on outcomes such as attendance, reading and math scores, and parent involvement.

Some foundations seek a broader perspective. Social Venture Partners in Seattle is developing a framework that will allow it to aggregate the outcome data provided by individual grantees. Outcome data will be grouped into clusters by topic (e.g., academic, after-school, capacity-building). Outcomes that are very specific to organizations and ventures, are reported anecdotally. The foundation recognizes that it cannot measure success "per share," as venture capitalists do, but it does not want to give up trying to measure outcomes.

REDF, which invests in non-profit enterprises employing low-income personnel, is in the process of establishing an OASIS system, a computer database which measures operational and social outcomes. REDF tracks financial information and 40 social outcomes plus outcomes customized to individual grantees' needs. Financial indicators include gross sales, gross profit, and net profit. Customized indicators may include inventory reliability, production wastage, and revenue per square foot. Social impact indicators include job retention, job placement, job promotion, wages, barriers to employment, reliance on public assistance, utilization of services, housing stability,

self-esteem, personal support, and involvement in the criminal justice system (Twersky, 2000). Currently, one grantee out of five uses the Oasis system. In its business plan, the organization specifies the outcomes to be achieved, and then a social outcomes expert works with the organization to develop an appropriate database to collect the necessary data. The database is then automated and becomes part of the Oasis system. REDF plans to incorporate indicators to measure the social returns on investment, but its framework is not yet operational (Twersky, 2000).

Evaluation Process

Both traditional and venture philanthropies evaluate virtually all grants, but grant makers differ in how they communicate expectations about evaluation, the measures they use, and the degree of foundation involvement. Expectations regarding reporting and evaluation are communicated by traditional foundations either in proposal guidelines, an application form, or the award letter. Either the foundation or an external consultant designs the evaluation. Large initiatives are usually evaluated more rigorously and employ an outside evaluator who may work with the grantee and foundation staff in designing the evaluation. Otherwise, the measures to be used are rarely developed by the grantee. Reporting is usually annual and includes financial information and a narrative. If resources permit, site visits may be made.

Venture philanthropies often require grantees to write a business plan that spells out specific milestones to be met and measures to be used. Foundation staff work actively with grantees in developing outcomes and remain active in monitoring performance and assisting organizational development. Typically, the business plan replaces the usual application form or proposal. It also defines the outcomes to be achieved. At the Edna McConnell Clark Foundation, for example, the grantee organization presents a business plan with milestones, and outcomes are negotiated among all parties. All grantees must establish milestones, indicators, a growth plan, resources, and plans for diversifying funding. As with most of the other venture philanthropies we studied, Clark portfolio managers are intimately involved in the movement toward milestones. The portfolio managers report on a weekly basis.

Capacity and Resources for Evaluation

Capacity. Almost all of the foundations provide grantees with technical assistance for developing evaluation plans. Sometimes assistance is on a one-to-one basis with a program officer or evaluation consultant, and sometimes, as in the case of Hewlett and its Bay Area School Reform Collaboration, the foundation takes an active role in study and research design.

Typically, two or three foundation staff members are involved in assessment. Their job is to help establish objectives and supervise outside evaluators. Of the 12 foundations we interviewed, only Pew had a large in-house evaluation team. Its internal evaluation unit consists of 11 people: a director, four program officers, three associates, and three support staff. Evaluation staff work with program staff to develop measurable objectives for each grant, develop evaluation designs, and hire and oversee external consultants. Pew reports, however, that it is moving toward using external consultants to better capture the outside perspective and to ensure credibility. At the Irvine Foundation, a small evaluation staff consisting of a director and assistant director focuses on assisting project officers rather than directly helping grantees.

For the venture philanthropies, assistance focuses on developing a business plan, milestones, and measures, as well as help in the actual establishment of a successful organizational operation. The venture foundations take an active role in all phases of a grantee organization.

Resources. The funding available for evaluation varies. Annual or semi-annual narrative and financial reports are generally expected conditions of a grant award. On the other end of the spectrum, Panasonic funds an on-site consultant at each school district with which it has a partnership. The consultants talk with district stakeholders and provide feedback about the progress of each partnership. They also conduct an annual site review based on a school effectiveness evaluation template. Venture philanthropies devote staff or consultant time to organize these evaluation measures outlined in business plans with the grantees.

Several foundations pay the costs of evaluation either by including evaluation provisions in the grant or hiring an evaluator themselves. The Robin Hood Foundation tells an independent evaluator what it is interested in learning. The evaluator specifies the data it needs, and the grantee provides the data. Each grantee is evaluated annually. Sometimes the outcomes are so straightforward—the number of meals served, for instance—that an evaluator is not needed. Robin Hood has also found that some grantees hire their own evaluators, usually because they want a more extensive evaluation than stipulated by the terms of the grant.

Provisions for evaluation can be made within a grant award. Typically the amount spent on evaluation varies by the size of the grant and the significance of the initiative. At the Colorado Trust, a minimum of 10 percent and usually much more (often 50 percent or more) of a grant goes to evaluation. For its After School Care Initiative, the Colorado Trust has included $150,000 per year for three years to be spent on evaluation at the 30 sites. When the Bay Area School Reform Collaborative originally received $25 million over five years from the Hewlett

Foundation, $1.5 million of that was set aside for evaluation. The grant was recently renewed and more money will be used for evaluation of the new phase.

Evaluation of Long-Term Effects

Several of the foundations we interviewed periodically survey a group of grantees to study the long-term effectiveness of their programs. Ashoka, for example, conducts periodic evaluations of its Fellows program to measure its effectiveness. A given cohort is surveyed several years after stipends have ended, and follow-up case studies are conducted with selected fellows. Ashoka's interests center on the impact of the grantee's work, and the impact the foundation has had in aiding its work. Echoing Green surveyed its last nine years of fellows (n = 220) to look at the social impact of their activities, and to inform fellows about their peers' work. Its survey focused on leadership development, organizational development, and social impact. However, the foundation has found it difficult to measure social impact; therefore, it requested input from fellows on how to define terms and how to use the measures.

Foundations may evaluate a cluster of grants in a given area. Pew recently reviewed a cluster of 25 grants made over a decade in a single program area to examine the impact and theory underlying a strategy. It brought in a team of external evaluators, and they focused on key questions such as the following (Rimel, undated):

1. Was the problem an appropriate one for the Trusts to have tackled?

2. Were some interventions and initiatives more effective than others?

3. Was the size of the investment right?

4. Did we help 'move the needle' on this?

Panasonic Foundation has partnerships with 13 school districts to which it provides ongoing technical assistance. Each partnership is based on a template of what constitutes effective schools. After about five years, the foundation conducts a formal assessment of the template to determine whether or not the foundation is adding value in the district. It focuses on developing the organizational and bureaucratic conditions for improved student learning. The foundation also tries to assess its own role as a partner—whether it is being strategic, the way it wants districts to be. Because Panasonic believes it can't claim a direct impact on student learning with its programs, it does not measure student outcomes.

Several foundations are in the process of establishing databases which can be used to examine long-term effectiveness. We have already discussed REDF's Oasis computer database for measuring operational and social outcomes. In addition, the Edna McConnell Clark Foundation is establishing a multi-purpose database to analyze its effectiveness; to use for institutional learning, developing business plans and milestones; to share information with other funders, practitioners, and policy makers; and to add information to the field of youth development. The Colorado Trust has found that grantees want to set-up a computerized data system for the After-School Care Initiative, which includes 30 sites. Once established, it will collect process and outcome data to track the extent to which the program is making an impact so that grantees can enlist other funders. The Trust has contracted with an evaluation firm to provide one-to-one technical assistance to each site. Each site will develop its own indicators.

Lessons Learned

In sum, our interviews provided the following generalizations about philanthropic foundations' approaches to program evaluation:

- Foundations use evaluation to judge their impact, assess their effectiveness in assisting grantees, and report to their board and the community at large. Generally, evaluation results are not related to continued funding.

- Many foundations are employing a business venture-capital model to organize and evaluate their grant giving. Evaluation often focuses on outcomes.

- Financial reports and activity narratives comprise much of the evaluation data. Typically, the larger the grant, the more rigorous the evaluation. Grantees may want rigorous outcome measures for themselves to judge their effectiveness and to assist in securing additional funding.

- Evaluation measures are usually tailored to the grant. Few foundations use common measures across all grants. However, some foundations are beginning to develop systems to track outcomes across grants and establish databases to share information among grantees and with others.

Appendix B:
Critical Review of Other Evaluation Methods

RAND investigated several current evaluation methods to determine whether they might be appropriate for the Broad Foundation to use in evaluating either individual grant programs or its portfolio of grants. The most notable of the current methods are the Balanced Scorecard method, cost-benefit analysis, and the Social Return on Investment method. Though each of these is a useful approach to evaluation, RAND found that none meets the needs of the Broad Foundation. Consequently, RAND developed a unique approach to foundation evaluation for Broad, one that suits the nature of programs focused on the development of school leaders.

For the purpose of general information, however, we include brief descriptions of these three methods below.

Balanced Scorecard

The Balanced Scorecard is a tool used by some non-profit organizations, such as the Massachusetts Special Olympics, and by some government organizations. Robert Kaplan et al. (1996) designed the Balanced Scorecard to help organizations monitor and evaluate their progress on non-financial measures. The scorecard is an assessment, strategic planning, and management tool that is used to identify organizational goals and translate goals into operational measures of performance. The approach is based on four processes: translating the vision, communication and alignment, business planning, and feedback and learning.

The scorecard is particularly useful because it requires that organizational effectiveness be viewed from four perspectives: the financial perspective, the internal business perspective, the innovation and learning perspective, and the customer perspective. The measures developed for each organization are unique since they are developed through a process that includes stakeholders of the organization. Many organizations have used the Balanced Scorecard successfully.

However, RAND does not recommend the Balanced Scorecard as the right evaluation tool for the Broad Foundation. First, the developers recommend that the scorecard be used at the level of the business unit, the unit which has responsibility for managing resources and producing the outcomes identified on the scorecard. But unlike a business unit, grantees rarely have control over the processes that produce the outcomes of interest. Indeed, many Broad-supported initiatives try to influence part of a larger system that supports schooling. Even when a grant

recipient is a school district with responsibility for the entire process, the Broad Foundation-funded initiative supports only a subset of measures that should appear on the district's scorecard. Second, much of the value of the Balanced Scorecard resides in the process of developing the scorecard itself. Performance measures are unique to each site and to the interest of the organization's stakeholders. Consequently, each Balanced Scorecard is unique. While such an organic process may develop buy-in and ensure stakeholder support for Balanced Scorecard measures, the process is not of particular benefit to the Broad Foundation, which is most interested in understanding the efficacy of its investments.

Cost-Benefit Analysis

Another widely-used tool for evaluating social programs is cost-benefit analysis. Cost-benefit analysis compares the costs of an entire program or the marginal cost of specific program characteristics to the value of the benefits generated by the program. As an evaluation tool, cost-benefit analysis helps organizations to measure the relative cost-effectiveness of programs. A challenge posed by benefit calculations is that a comprehensive cost-benefit evaluation typically compares the costs borne in one time period to the stream of benefits realized in future years. For instance, the cost-benefit analysis in Karoly et al. (1998) considered benefits realized up to 15 years later for one type of early intervention program and up to 30 years later for another. The benefits incorporated into that analysis included increased earnings, reduced rates of public assistance utilization, lower rates of criminal activity, and others.

However, this approach has limitations for use as an evaluation tool for venture philanthropy and specifically for a foundation like Broad. For ventures like those supported by the Broad Foundation, cost-benefit analysis becomes complex because it attempts to measure social costs and benefits that are not easily monetized and that are difficult to attribute to specific programs or activities. In addition, comparing the costs of a program to the benefits realized only a year or two after educational leadership development is unlikely to produce any meaningful monetized benefits. However, limited information about the likely benefits of a program may be elicited by measuring short-term outcomes and comparing them to other studies that have measured both short-term and long-term benefits. In such cases, a definitive comparison of costs and benefits is not feasible, but it may be possible to say, for example, that the early findings compare favorably or unfavorably to outcomes from studies that followed participants for longer periods of time. This analysis would require the computation of statistical estimation errors (as in Karoly et al., 1998).

A second difficulty is that the typical focus of cost-benefit analysis does not match that of most program evaluations. In cost-benefit analysis, the analysis typically focuses on the benefits the program generates for members of society beyond program participants; for example, these would include benefits to students in the form of increased earnings due to high school graduation, taxpayers in the form of reduced future taxes, or benefits to other members of society in the form of lower property loss from criminal activity. In other words, the analysis often assesses whether the benefits to the taxpaying public generated by the program outweigh the costs in terms of public funds used for the program (see discussions of cost-benefit analysis in RAND publication Karoly et al., 1998). In contrast, most program evaluations concentrate on the benefits a program accrues to program participants.

Social Return on Investment

The Roberts Enterprise Development Fund has been supporting the development of a measure of the "Social Return on Investment," a specialized method of cost-benefit analysis. The SROI is a performance measure of the social impact of investments used to determine their effectiveness. SROI is intended to document the effort-cost savings "for social sector managers to use in advocating for financial support of their work" (Roberts Enterprise Development Fund, 8). Unlike the type of comprehensive cost-benefit analysis described earlier, SROI defines "benefits" more narrowly as the "various cost savings, reductions in spending and related benefits that accrue [directly to participants] as a result of that social service activity" (Emerson, Wachowicz and Chun, 2000, 139). A major giving area for REDF is job training, and the costs savings or benefits attributed to the program are public assistance foregone by employed participants (e.g., food stamps, TANF) and tax revenues generated by employed participants' wages. However, the calculation does not consider more far-reaching costs or benefits that accrue to other parties, such as lower property loss from criminal activity.

SROI is an interesting model and may become more applicable to Broad programs as it becomes more fully developed. However, its applicability to Broad Foundation evaluation efforts is limited at this time. First, the Broad Foundation invests in educational leadership development, and the benefits to participants are not the outcomes of interest to the foundation. Second, SROI is an expensive system to implement. As of 1999, REDF, in collaboration with partners, spent over $1.3 million developing the system, which is now being piloted with 11 REDF grants. Furthermore, because the outcome measures are customized to each project, no efficiencies arise from using the model. Most important, perhaps, is that REDF cautions against comparing rates of return among projects or calculating an aggregate rate of return across projects (Emerson,

Wachowicz and Chun, 2000, 157), severely limiting its use as a tool to assess the Broad Foundation's portfolio.

Appendix C:

Sample Logic Models

RAND developed a set of hypothetical logic models as examples of how Broad-funded programs might be diagrammed. They illustrate how the logic model technique may be applied to programs targeted at school governance, school leaders' management skills, and labor negotiations between teachers and school administrators.

All of the models use an "activities" organizational structure, though they range in complexity from relatively simple principal recognition programs to more complicated performance pay initiatives and collective bargaining systems. Each model includes boxes arranged chronologically from left to right representing a program's main tasks; the arrows indicate relationship and sequence. The bullet points below the boxes list program indicators available for evaluation purposes at each stage of the program.

These models are intended simply as examples. Grantees and Broad Foundation staff will need to develop unique logic models for each program as they begin to determine the requirements of a grant and the type of evaluation that a grant will undertake. The *W. K. Kellogg Foundation Evaluation Handbook* describes alternative organizational structures for logic models, including models based on program outcomes, program theory, and a combination approach (W. K. Kellogg, 1998, 36-42). These alternative structures may prove useful in some instances. RAND developed samples of activities logic models because this structure is the most intuitive of the various organizational strategies and may be easiest for grantees to create.

Four sample logic models follow in Figures C.1-4.

Program description: An organization sponsors a regional or national prize program to recognize successful principals.

Identify criteria for prizes	Advertise prize	Collect nominations	Select prize winners	Announce/ publicize winners	Disseminate winners' stories

Advertise prize
- Public awareness of prize
- Educators' awareness of prize
- Media coverage of prize (e.g., # of TV spots, # of articles, # of radio spots)

Collect nominations
- # of nominations
- Diversity of nominees (e.g., geography, position, accomplishments)
- "Quality" of candidates (e.g., credentials)

Select prize winners
- # of winners
- Accomplishments of winners

Announce/publicize winners
- Media coverage
- Public awareness
- Educator awareness

Disseminate winners' stories
- # of requests for information
- Media coverage
- Public awareness
- Educator awareness

Have winners provide technical assistance to others
- # of requests
- # of districts and schools served
- # hours served
- Types of assistance provided
- Impact of technical assistance

Award winners cash prizes
- $ awarded

Winners use cash prizes
- Uses of funds
- Impact of funds (personal, school outcomes)

Figure C.1. Logic Model for Program to Recognize Successful Principals

Program description: An organization offers training to potential school board members to improve governance in targeted districts.

Identify civic-minded leaders within targeted districts					
Ask civic-minded leaders to nominate individuals for training program	Invite nominees to participate in training program	Conduct training program on education issues, school board governance, & campaigning	Leaders participate in K-12 public education	Leaders run for school board; candidate pool increased	Improved school board governance

Invite nominees to participate in training program
- # invited
- % yield
- # school board seats represented
- # districts represented
- "Quality" of nominees/class

Conduct training program on education issues, school board governance, & campaigning
- Hours of training
- # of workshops
- List of workshop topics
- Change in participants' knowledge and interest in K-12 public education
- Participant's reaction to training (smile sheets)

Leaders participate in K-12 public education
- Pre/post inventory of knowledge about education issues
- # who participate in education leadership roles (e.g., school councils, business-education partnerships, advocacy organizations, $ donated, hours volunteered at school)

Leaders run for school board; candidate pool increased
- # who register to run as candidates
- Competitiveness of election (e.g., number of candidates, election margins)
- Funds raised
- # and type of endorsements

Improved school board governance
- # trained individuals elected
- Proportion of board trained by program
- Change in policy directions of board
- Change in climate of school district

Figure C.2. Logic Model for School Board Recruitment and Training Program

Program description: A district adopts a performance pay program and implements it by setting criteria for awards and opportunities to build teachers' skills and knowledge.

Committee develops performance pay criteria based on knowledge and skills

School district and union publicize performance pay criteria

- Teacher awareness/ support
- Public awareness/ support
- # who 'register' for performance pay

Committee develops group incentive pay criteria

School develops plan to increase school performance

- # of schools developing plans
- "Quality" of plans
- Schoolwide belief that performance can improve
- Schoowide belief that performance will be rewarded

District and union offer opportuni-ties for teachers to build capacity

- # of courses/ workshops
- # of courses/ workshops related to criteria
- # of courses/ workshops that meet criteria of "effective" PD
- % of teachers in school participating

Teachers participate in capacity-building activities

- Hours of training
- # of classes
- Participant reactions
- Change in participant knowledge of content or pedagogy

Peers/administrators evaluate individual teachers' use of knowledge/skill criteria

Teachers change instruction

- Change in curriculum
- Change in pedagogy
- Change in choice of learning activities
- Change in materials
- Change in test preparation
- Change in allocation of instructional time

Student learning increases

- Test scores
- Student grades
- Pace/ coverage of material
- # of students on grade level (measured by in-class assessments)

Schoolwide performance improves

- Test scores
- Dropout rates
- Retention rates
- Grades (schoolwide)
- Attendance
- Transition rates (success of transition)

District awards performance pay to individuals

- $ awarded
- Average awarded
- Reports of relationship of incentives to behaviors
- # affected

School/ teachers awarded group incentive pay

- $ awarded
- Average awarded
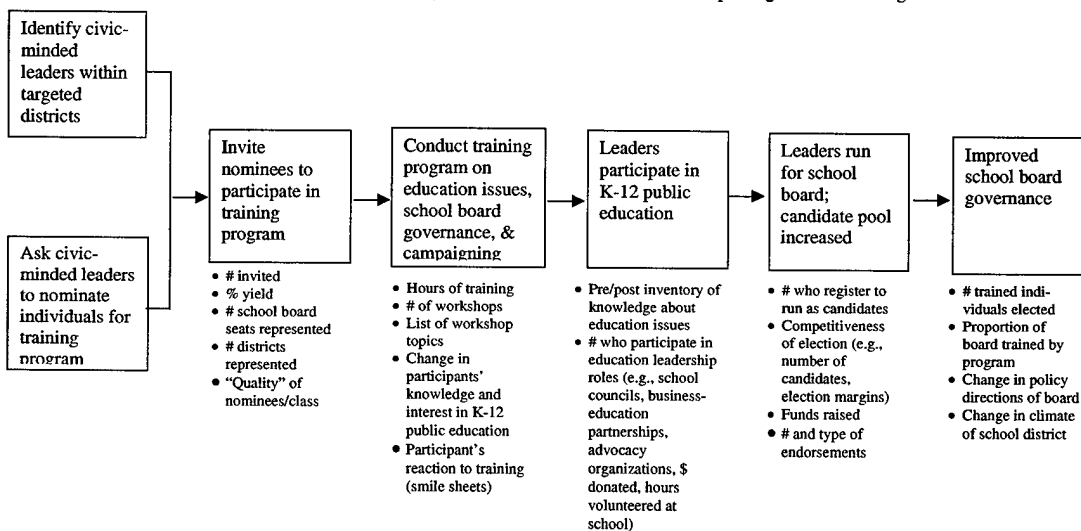- Reports of relationship between incentives and behaviors
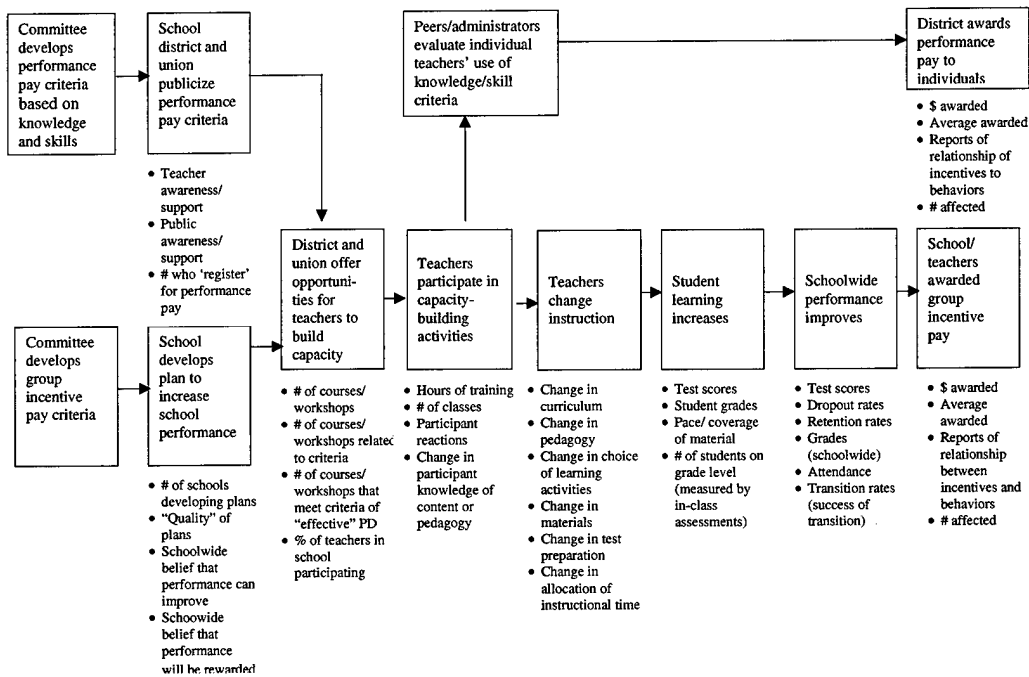- # affected

Figure C.3. Logic Model of Performance Pay Program

Program description: A program trains and guides districts in collective bargaining in order to bring quality and accountability issues into contract negotiation.

Identify districts with expiring contracts

Research terms of existing teachers' contracts

Approach select districts and stake-holders about innovative bargaining

- # approached
- % who participate

Project offier studies "innovative" collective bargaining agreements

Identify individuals who can serve as "innovative" bargaining coaches

Project officer trains coaches

- # trained
- Knowledge/skills of coaches

Conduct "policy game" around collective bargaining

Develop model of collective bargaining agreements

- # of districts and union leaders who attend
- attendees' response to game

Develop consensus around collective bargaining principles

Publicize materials for interested parties

- # of requests for materials/coaches
- # of hits on web sites

Develop and ratify new teacher contracts

- Pre/post attitudes about bargaining process
- Terms of contracts which reflect "new unionism"
- Speed at which agreements are reached
- % of time agreements reached in designated negotiation period
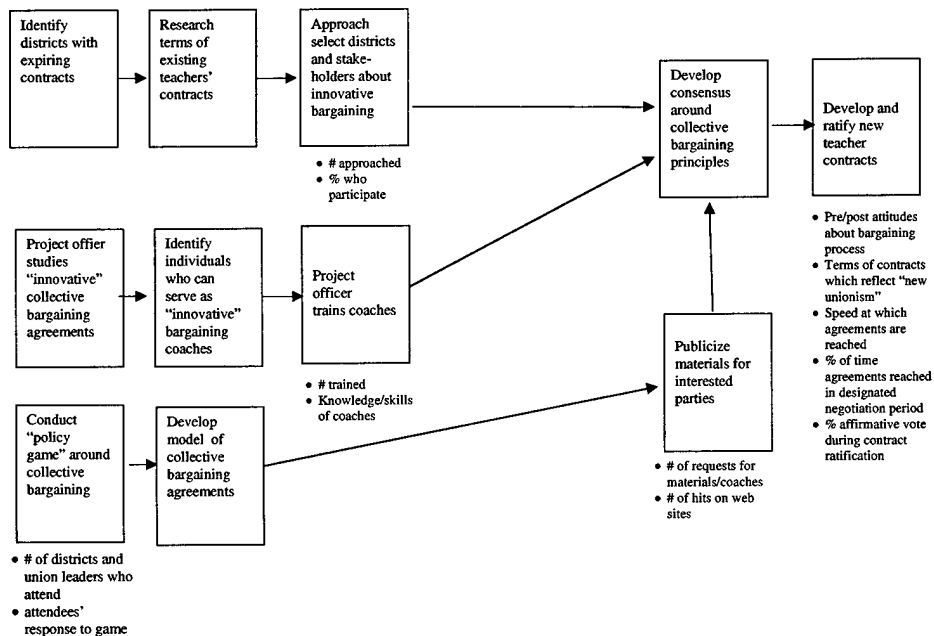- % affirmative vote during contract ratification

Figure C.4. Logic Model for Program to Reform Labor Relations